

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

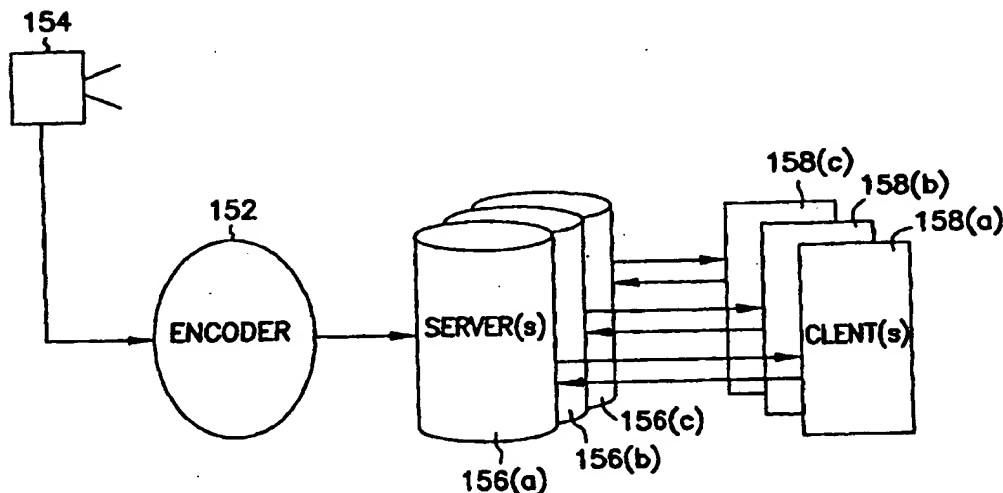
**THIS PAGE BLANK (USPTO)**



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>H04N 7/173</b>		A1	(11) International Publication Number: <b>WO 98/34405</b>
			(43) International Publication Date: <b>6 August 1998 (06.08.98)</b>
(21) International Application Number: <b>PCT/US98/01810</b>		(81) Designated States: JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: <b>30 January 1998 (30.01.98)</b>		Published With international search report.	
(30) Priority Data:			
60/036,661	30 January 1997 (30.01.97)	US	
60/036,662	30 January 1997 (30.01.97)	US	
08/819,586	14 March 1997 (14.03.97)	US	
(71) Applicant: VXTREME, INC. [US/US]; 675 Almanor Avenue, Sunnyvale, CA 94086 (US).			
(72) Inventors: CANNON, Anthony; 1600 Villa Street #334, Mountain View, CA 94041 (US). DEL VAL, David; 1012 Bonita Avenue, Mountain View, CA 94040 (US). KLEMETS, Anders; 655 S. Fair Oaks Avenue #K-313, Sunnyvale, CA 94086 (US).			
(74) Agent: VIKSNINS, Ann, S.; Schwegman, Lundberg, Woessner & Kluth, P.O. Box 2938, Minneapolis, MN 55402 (US).			

(54) Title: VCR-LIKE FUNCTIONS RENDERING VIDEO ON DEMAND



## (57) Abstract

A method for displaying streamed digital video data on a client computer. The client computer is configured to receive the streamed digital video data from a server computer via a computer network. The streamed digital video data is transmitted from the server computer to the client computer as a stream of video frames. The method includes receiving a first plurality of video frames at the client computer. The plurality of video frames represents a subset of the stream of video frames. The stream of video frames comprises independent playable video frames and dependent playable video frames. The method further includes displaying the first plurality of video frames on a video display terminal associated with the client computer. There is further included issuing a rewind command from the client computer to the server. The rewind command causes a second plurality of video frames of the stream of video frames different from the first plurality of video frames to be streamed from the server computer to the client computer. The second plurality of video frames has been streamed at least once to the client computer.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## VCR-LIKE FUNCTIONS RENDERING VIDEO ON DEMAND

5

## BACKGROUND OF THE INVENTION

10

The present invention relates to improved techniques for displaying video images transmitted over a computer network. More particularly, the present invention relates to improved methods and apparatus for implementing control features, such as play, rewind, fast forward, pause, stop, record, and the like, on a real-time video stream and/or live video stream delivered via a computer network from server computers to client computers.

It is well known that digital video data may be manipulated and rendered using computers. In a computer network, e.g., a client-server computer network, one or more computers may be employed to receive the analog video data (e.g., from a video camera), encode that analog video data to a suitable digital format, and store the digital video data file. Using a computer coupled to the network, a user may, at a subsequent time, request the pre-stored digital video data file for display on a video display associated with the client computer.

As computers become more pervasive in the workplace and in the home, the demand for digital video services correspondingly increases. By way of example, it has been recognized that it is possible to employ networked computers as a mass communication medium whereby a pre-stored digital video file may be transmitted from a server to one or more client computers to permit the client computers to display the images after the file is received. This technology may be employed to, for example, deliver movie or training video clips from a central server to one or more client computers for display.

In the above example, it is typically necessary for the client computer to receive the entire pre-stored digital video file prior to rendering the images. Real-time video

streaming, on the other hand, refers to the situation wherein the client computer renders the images while they are streamed from the server computer. In some applications, real-time video streaming is favored since it permits the user to begin viewing video frames shortly after requesting the video file instead of having to wait until the entire pre-stored file is downloaded from the server computer.

It has been found, however, that real-time video streaming is difficult to implement on heterogeneous, lossy networks such as corporate intranets or the Internet, i.e., the well-known international computer network that links, among others, various military, governmental, educational, nonprofit, industrial and financial institutions, commercial enterprises, and individuals. This is because real-time digital video applications, as are all digital video applications, are resource-intensive to implement. Even with compression, the transmission of quality video clips (i.e., those with acceptable frame rate and frame quality) places a heavy bandwidth burden on the computer network. For that reason, real-time video streaming has traditionally been implemented on proprietary and expensive networks that are capable of supporting a high bit rate (e.g., private high-speed local area networks (LAN) or dedicated data links).

Furthermore, real-time video data is time-sensitive, i.e., the data packets containing the real-time video data must be received in a timely manner in the proper sequence for acceptable display. In bandwidth limited networks, e.g., corporate intranets which support a high number of users or heterogeneous, lossy public networks such as the aforementioned Internet, the time-sensitive nature of real-time digital video data poses special challenges. There is, for example, less time to retransmit a lost data packet because if the time for displaying a given data packet at the client computer has passed, there is little use for that data packet if and when it arrives.

It has also been found that real-time digital video streaming poses complex frame synchronization issues. Since the video frames to be displayed are not stored with the client computer, there is no pre-stored file on which to perform control features such as rewind, fast forward, play, and pause. Typically, the video frames necessary for performing these functions are requested from the server computer itself. Responsive to the control commands, the video frames necessary for performing the requested control feature are then streamed to the client computer, typically over the same data connection to minimize the latency associated with opening another data connection. As can be appreciated by those skilled, complications can arise while switching among different groups of video frames, some of which may be waiting to be sent at the server, waiting to be displayed at the client, in transit through the network, or lost.

The complexity involved in implementing control features on real-time video stream is further compounded by the requirement of low latency, which is imposed by real-time video applications, i.e., the requirement that any delay between the time a given video frame is transmitted from the server computer and the moment it is rendered at the client computer be minimized. Unless these control features are properly implemented, undue latency may occur and/or the quality of the viewing experience may degrade.

All the challenges discussed above also apply to live video streaming. In live video streaming, the video data may be digitized in one location, then encoded and transmitted to a client computer (via a server) substantially instantaneously (with some delay due to the encoding and transmission of video frames) for display. Live video and real-time video may be thought of as subsets of streamed video since both live video frames and real-time video frames are rendered as they are streamed from the server. Live video rendering, however, results in the display of most recently encoded video frames whereas real-time video rendering may result in displaying past video frames of a currently recorded event or even an event that happened and was recorded a long time ago. As can be appreciated, live video streaming applications are even more sensitive with respect to the data packets transmitted via the network. This is because the live event being recorded continues to unfold, and video frames related thereto continue to be formed and require to be displayed as time goes by.

Frame synchronization issues pertaining to live video streaming are further complicated by the fact that the digital video data file at the server is being formed at the same time the video frames pertaining thereto are displayed at the client computer. This is because copies of video frames sent to the client computer are also stored in the server in a digital video data file for subsequent use. Accordingly, there are complexities involved when, for example, the user wishes to switch from a live video viewing mode to a rewind mode, view past video frames for a few seconds, and fast forward to the end of the still growing digital video data file to again play live video frames. Because of the complexities involved, as well as the bandwidth and latency requirements, prior art attempts at implementing control features on live video streams have largely been unsatisfactory. While this is true for most networks, it is particularly true for the Internet wherein the transport network is typically lossy and outside the control of the owner of the server and/or the client computer, and wherein the bandwidth available is both limited and subject to fluctuations.

In view of the foregoing, there are desired improved methods and apparatus for implementing control features on real-time video streams and/or live video streams transmitted via a computer network from server computer(s) to client computer(s).

## SUMMARY OF THE INVENTION

The invention relates, in one embodiment, to a method for transmitting streamed digital video data from a server. The server is configured for coupling to a client  
5 computer via a computer network. The method includes inputting a first plurality of data packets into a server play-out buffer of the server. The first plurality of the data packets contains video frames representing the streamed digital video data. An output of the server play-out buffer is configured to be coupled to a network data connection for transmitting the first plurality of the data packets to the client computer.

10 The method includes receiving, using a retransmission buffer, the first plurality of the data packets from the server play-out buffer. An output of the retransmission buffer is coupled to the network data connection. The method further includes outputting the first plurality of the data packets from the server play-out buffer onto the network data connection for transmitting the data packets to the client computer via the  
15 computer network.

In another embodiment, the invention relates to a method for displaying streamed digital video data on a client computer. The client computer is configured to receive the streamed digital video data from a server computer via a computer network. The streamed digital video data is transmitted from the server computer to the client  
20 computer as a stream of video frames. The method includes receiving a first plurality of video frames at the client computer. The plurality of video frames represents a subset of the stream of video frames. The stream of video frames comprises independent playable video frames and dependent playable video frames.

The method further includes displaying the first plurality of video frames on a  
25 video display terminal associated with the client computer. There is further included issuing a rewind command from the client computer to the server. The rewind command causes a second plurality of video frames of the stream of video frames different from the first plurality of video frames to be streamed from the server computer to the client computer. The second plurality of video frames has been  
30 streamed at least once to the client computer.

In yet another embodiment, the invention relates to a computer readable medium containing computer-readable instructions for implementing control features configured for controlling a display of streamed digital video data at a client computer as the client computer transitions from a first control mode to a second control mode. The client



computer is configured for coupling to a server computer via a computer network. The streamed digital video data is transmitted from the server computer to the client computer as a stream of video frames comprising independent video frames and dependent video frames. The stream of video frames is encapsulated in a plurality of data packets each having a unique packet sequence number and a unique timestamp.

5 The computer readable instructions include computer readable instructions for sending a control command and a time parameter from the client computer to the server computer. The control command represents a command to the server to transmit a first plurality of video frames of the stream of video frames to the client computer in accordance with the

10 second control mode.

The computer readable instructions also include computer readable instructions for ascertaining, responsive to the control command and using the server computer, a first independent video frame to transmit to the client computer. The first independent video frame is selected responsive to the time parameter. The computer readable

15 instructions further include computer readable instructions for transmitting from the server computer to the client computer a packet sequence number associated with the first independent video frame. The computer readable instructions further include computer readable instructions for streaming the first plurality of video frames of the stream of video frames starting from the first independent video frame from the server

20 computer to the client computer to permit the first plurality of video frames to be displayed at the client computer.

These and other features of the present invention will be described in more detail below in the detailed description of the invention and in conjunction with the following figures.

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1A depicts in accordance with one embodiment of the present invention a computer network 1 suitable for implementing the inventive streamed video display technique.

Fig. 1B illustrates an embodiment of the invention wherein an encoder furnishes video data from a video source to multiple servers.

Fig. 2 is a block diagram of an exemplar digital computer, representing a computer suitable for implementing either the server computer or the client computer of the present invention.

Fig. 3 illustrates, in accordance with one embodiment of the invention, a VXF-formatted file, representing a file suitable for streaming encoded video data from the source, e.g., a video camera, to the server and the client application.

Fig. 4 depicts, in accordance with one embodiment of the invention, two video streams: a play stream and a fast-forward stream to facilitate discussion.

Fig. 5 illustrates, in accordance with one embodiment of the invention, a simplified flowchart illustrating the implementation of certain control features such as play and fast forward.

Fig. 6 illustrates the steps involved, in one embodiment of the present invention, to implement the step for ascertaining the first data packet for sending for the real-time play mode.

Fig. 7 illustrates the steps involved in implementing the step for ascertaining the first data packet for sending for the live play mode, in accordance with one embodiment of the present invention.

Fig. 8 illustrates the steps involved, in one embodiment of the present invention, to implement the step for ascertaining the first data packet for sending when transitioning from the real-time play mode (or other modes except live play) to the fast forward mode.

Fig. 9 illustrates the steps involved, in one embodiment of the present invention, to implement the step for ascertaining the first data packet for sending when transitioning from the fast forward mode to the real-time play mode.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention will now be described in detail with reference to a few preferred embodiments thereof as illustrated in the accompanying drawings. In the following description, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art, that the present invention may be practiced without some or all of these specific details. In other instances, well known process steps have not been described in detail in order to not unnecessarily obscure the present invention.

In accordance with one aspect of the present invention, there are provided improved techniques for streaming real-time video data from the server computer to the client computer for display. In accordance with this aspect of the present invention, digital data pertaining to the real-time video stream is transmitted from the server computer to the client computer as data packets in either one of two streams: a play stream and a fast forward stream. As will be explored in great details herein, the use of a separate fast forward stream of video frames advantageously permits the fast forward feature to be implemented with lower bandwidth requirements and improved display quality. As the user switches from the play mode to the fast forward mode and vice versa, data packets containing video data in either the play stream or the fast forward stream are transmitted from the server computer to the client computer for display.

The play stream includes frames which, when played at the designated frame rate by the renderer application in the client computer (about 10 frames per second in one example), maximizes quality while minimizing the bandwidth requirement. This normal play mode represents the mode that the user normally employs to view the video frames (e.g., watching a movie or a video clip). It should be understood that normal play is typically accompanied by sound (typically furnished to the client via a data connection different from the one used to transmit video), and perhaps other applets. The fast forward stream includes frames which, when played at the designated frame rate by the renderer application in the client computer, provides the user with a fast forward effect while advantageously keeping the display quality higher and the bit rate lower than would have been possible had the play stream been employed for fast forwarding. This aspect of the invention is discussed in detail later herein. By way of example, the fast forward stream may be played at 5 frames per second, which displays frame at five times the play speed.

In accordance with another aspect of the present invention, the data packets traverse at least two buffers prior to arriving at the render application in the client computer for display: a retransmit buffer at the server computer and a client play-out buffer at the client computer. In one embodiment, a server play-out buffer is provided at the server as well to facilitate efficient data packet transmission. The use of the client play-out buffer and/or the server play-out buffer advantageously maintain(s) small supply of data packets available at the client ready for display, thereby minimizing impacts on the viewing experience due to temporary fluctuations in the network's available bandwidth and the temporary disruptions to the transmission of data packets through the computer network, e.g., due to temporary network congestion.

In accordance with yet another aspect of the present invention, there are provided novel and efficient implementations of control features, such as play, rewind, fast forward, pause, stop, record, and/or the like. In one embodiment, the control features are implemented to maximize the user's familiarity with common video cassette recorder (VCR) control features. Using the control features of the present invention, the user is able to control, in a user-friendly and intuitive manner, the transmission and display of the video frames pertaining to a real-time video stream or, advantageously, even a live video stream at the client computer. This aspect of the invention is particularly advantageous since the user may, using the inventive technique, flexibly control the display of streamed real-time video frames without being unduly constrained by the real-time nature of the data or the inherent time-sensitive nature of the transmitted real-time data packets. In one embodiment, the transition between the different control modes, e.g., from play to fast forward, from rewind to play, and the like, are made efficient and synchronized, thereby maximizing the display quality of video frames.

To further facilitate discussion of the foregoing, Fig. 1A depicts in accordance with one embodiment of the present invention a computer network 100, representing a computer network suitable for implementing the inventive real-time video display technique. Computer network 100 may, for example, represent a portion of the aforementioned Internet or corporate intranet.

In Fig. 1A, computer network 100 includes a server computer 102 and a client computer 104. There is also shown a video camera 106. In the present example, video camera 106 represents the device for recording video data. The recorded video data may then be digitized and encoded by encoder 110 into the proper digital data format for transmission to either server 102 or memory 115 for storage. Encoder 110 represents, in one embodiment of the invention, the video source from which data may be streamed to the client via the server. Encoder 110, which may be implemented in hardware or software, may also perform compression on the raw digital video data to improve

storage and transmission efficiency. One suitable encoding scheme is disclosed in a commonly assigned co-pending patent U.S. Patent Application Serial Number 08/623,299, filed March 28, 1996, incorporated herein by reference for all purposes.

5 Data packets outputted by encoder 110 (or retrieved from memory 115) are then buffered within server play-out buffer 112 for transmission to client computer 104. Although memory 115 is depicted as nonvolatile disk memory in Fig. 1A, it may represent any suitable type of memory device, including volatile semiconductor memory. As will be discussed earlier, the file of data packets stored within memory 115 may be employed by client computer 104 to facilitate rewind, fast forward, and  
10 other control modes.

As each data packet or group of data packets is outputted from server play-out buffer 112 onto data connection 114 for transmission (e.g., responsive to a command from client computer 104 which is received by server computer 102 via a control connection 116), the same data packet or group of data packets is input into retransmit  
15 buffer 118 at the server. Control connection 116 and data connection 114 have been discussed in detail in a commonly assigned co-pending U.S. Patent Application Serial No. \_\_\_\_\_ (Atty Docket No.: VXTMP001/VXT704), entitled "Methods and Apparatus for Implementing Control Functions in a Streamed Video Display System."

Retransmit buffer 118 represents in one embodiment a first-in-first-out (FIFO)  
20 buffer which retains for a limited time a data packet transmitted from server play-out buffer 112. As new data packets are input into retransmit buffer 118, old data packets (starting with the oldest data packets) are discarded from transmit buffer 118. The use of the retransmit buffer advantageously facilitates the rapid retransmission of a data packet therein if that data packet is requested by client computer 104 for retransmission  
25 (e.g., in the event a data packet is detected to be missing by client computer 104). Retransmit buffer 118 is preferably sized such that a data packet stays in retransmit buffer 118 slightly longer than the average latency period between the time a data packet is transmitted from server 102 and when it is displayed at client computer 104. There is no need, in one embodiment of the invention, for the retransmit buffer 118 to be much  
30 larger than mentioned since, due to the time-sensitive nature of real-time video and/or live video, it is typically not useful to keep a data packet therein long past the time it is required at client computer 104 for display.

As data packets are received by client computer 104 from data connection 114, they are inputted into client play-out buffer 120 to be displayed by renderer application  
35 122. Client play-out buffer 120 may represent, in one embodiment, a FIFO buffer. Client play-out buffer 120 and/or server play-out buffer 112 are typically sized

appropriately to minimize latency while taking into account the reliability and stability of network 100 through which data connection 114 traverses. If the network is fairly reliable and its bandwidth is fairly stable, client play-out buffer 120 and/or server play-out buffer 112 may be relatively small to minimize the latency between the time a data packet is outputted by encoder 110 and the time it is displayed at the client computer. On the other hand, a larger client play-out buffer 120 and/or server play-out buffer 112 may be able to more effectively insulate renderer application 122 from temporary bandwidth capacity fluctuations and disruptions in the transmission of data packets from server computer 102.

Client play-out buffer 120 may be, but is not required to be, equal in size to retransmit buffer 118 since retransmission of a data packet from retransmit buffer 118 is typically useful only prior to its turn to be displayed, i.e., prior to the time that data packet is needed at the output of client play-out buffer 120. It should be appreciated that the buffers shown herein may be implemented, in one embodiment, in software and may be dynamically sized as needed. Renderer application 122 may, in this example, include a decoder for decoding the encoded digital data.

Although only one control connection 116 and one control connection 114 are shown in Fig. 1A to facilitate ease of discussion, a given real-time video session may involve multiple data and control connections for the multiple data streams, e.g., video, audio, annotations, and the like. Further, encoder 110 may be physically implemented on a computer other than server 102. In one embodiment, a single encoder may furnish video data to multiple servers, possibly located at geographically dispersed locations, to permit those servers to service requests from any number of client computers. Fig. 1B illustrates this embodiment wherein encoder 152 furnishes video data from video camera 154 to multiple servers 156. Using a suitable protocol, clients 158 may request video data packets from one or more of servers 156 to display video frames at clients 158.

In one particularly advantageous embodiment, client 158 represents a plug-in software application that plugs into a web browser executing on a client computer. When client application 158 is initiated, it establishes a control connection with one or more servers 156 and requests therefrom the desired video file(s). If the desired video file(s) could be found on server 156, server 156 may begin streaming video data packets, either from memory 115 or directly from encoder 110, to the requesting client application 158 via a data connection. Via the control connection, the user may also issue control commands, e.g., fast forward, rewind, pause, unpause, record, and the like, which alter the order in which video data packets is delivered to client application 158.

The servers and client computers of the present invention may be implemented using any number of well known computers capable of being interconnected via a computer network. In some cases, client computer 104 may represent a special purpose device for displaying data, e.g., a digital television set having therein the decoder and the appropriate renderer software or hardware. Fig. 2 is a block diagram of an exemplar digital computer 202, representing a computer suitable for implementing either server computer 102 or client computer 104 of Fig. 1. Digital computer 202 may include a display screen (or monitor) 204, a printer 206, a floppy disk drive 208, a hard disk drive 210, a network interface 212, and a keyboard 214. The digital computer 202 includes a microprocessor 216, a memory bus 218, random access memory (RAM) 220, read only memory (ROM) 222, a peripheral bus 224, and a keyboard controller 226. The digital computer 202 can be a personal computer (such as an Apple computer, e.g., an Apple Macintosh, an IBM personal computer, or one of the compatibles thereof), a workstation computer (such as a Sun Microsystems or Hewlett-Packard workstation), or some other type of computer.

The microprocessor 216 is a general purpose digital processor which controls the operation of the digital computer 202. The microprocessor 216 can be a single-chip processor or can be implemented with multiple components. Using instructions retrieved from memory, the microprocessor 216 controls the reception and manipulation of input data and the output and display of data on output devices.

The memory bus 218 is used by the microprocessor 216 to access the RAM 220 and the ROM 222. The RAM 220 is used by the microprocessor 216 as a general storage area and as scratch-pad memory, and can also be used to store input data and processed data. The ROM 222 can be used to store instructions or program code followed by the microprocessor 216 as well as other data.

The peripheral bus 224 is used to access the input, output, and storage devices used by the digital computer 202. In the described embodiment, these devices include the display screen 204, the printer device 206, the floppy disk drive 208, the hard disk drive 210, and the network interface 212, which is employed to connect digital computer 202 to the network. The keyboard controller 226 is used to receive input from keyboard 214 and send decoded symbols for each pressed key to microprocessor 216 over bus 228.

The display screen 204 is an output device that displays images of data provided by the microprocessor 216 via the peripheral bus 224 or provided by other components in the digital computer 202. The printer device 206 when operating as a printer provides an image on a sheet of paper or a similar surface. Other output devices such

as a plotter, typesetter, etc. can be used in place of, or in addition to, the printer device 206.

The floppy disk drive 208 and the hard disk drive 210 can be used to store various types of data. The floppy disk drive 208 facilitates transporting such data to other computer systems, and hard disk drive 210 permits fast access to large amounts of stored data.

The microprocessor 216 together with an operating system operate to execute computer code and produce and use data. The computer code and data may reside on the RAM 220, the ROM 222, the hard disk drive 220, or even on another computer on the network. The computer code and data could also reside on a removable program medium and loaded or installed onto the digital computer 202 when needed. Removable program mediums include, for example, CD-ROM, PC-CARD, floppy disk and magnetic tape.

The network interface circuit 212 is used to send and receive data over a network connected to other computer systems. An interface card or similar device and appropriate software implemented by the microprocessor 216 can be used to connect the digital computer 202 to an existing network and transfer data according to standard protocols.

The keyboard 214 is used by a user to input commands and other instructions to the digital computer 202. Other types of user input devices can also be used in conjunction with the present invention. For example, pointing devices such as a computer mouse, a track ball, a stylus, or a tablet can be used to manipulate a pointer on a screen of a general-purpose computer.

The invention can also be embodied as computer readable code on a computer readable medium. The computer readable medium is any data storage device that can store data which can be thereafter be read by a computer system. Examples of the computer readable medium include read-only memory, random-access memory, CD-ROMs, magnetic tape, optical data storage devices. The computer readable code can also be distributed over a network coupled computer systems so that the computer readable code is stored and executed in a distributed fashion.

In accordance with one aspect of the present invention, the encoder encodes video data in the VXF format for streaming to the client computers (via servers). Fig. 3 illustrates, in accordance with one embodiment of the invention, a VXF-formatted file, representing a file for streaming encoded video data from the source, e.g., a video camera, to the server and the client application.



In general, a VXF-formatted file 300 includes a file header 302, which may include data such as the length of file 300, the type of compression employed (e.g., HVQ), and other bibliographic information pertaining to the file. VXF-formatted file 300 further includes a compression header 304, which may contain data regarding the parameters employed in encoding file 300. The data in compression header 304 may be employed by a decoder at the client computer, for example, in order to decode the data therein.

Actual encoded video data is stored in a plurality of data packets 306. Each of data packets 306 generally includes a packet header 308 and a payload 310. Payload 310 may be employed to transmit any type of data. In the context of the present disclosure, payload 310 is employed to transmit encoded video, audio, or annotation data to the servers and/or the client computers. Packet header may also include data that indicates whether the data packet is associated with the play stream or the fast forward stream. The advantages associated with providing both a play stream and a fast forward stream will be discussed later herein.

In one embodiment, packet header 310 also includes data which indicates whether this data packet is associated with a video frame that is playable, seekable, or fast-forwardable. A video frame that is playable represents a video frame suitable for display while the client application is in either the real-time play mode or the live play mode. A video frame that is seekable permits the server to skip in reverse or forward through groups of frames as the server seeks among the seekable frames. A seekable frame is preferably a frame that is decodable without depending on information from its neighboring frames. A video frame that is fast-forwardable represents a video frame suitable for display while the client application is in the fast forward mode. This aspect of the invention is also discussed in detail later in connection with subsequent Fig. 4.

Packet header 310 may also include a packet number, which is assigned sequentially to each of the data packets. By monitoring the sequence of packet numbers associated with the received data packets, a client application may be able to ascertain whether a given data packet is missing (e.g., due to network problem) and to request retransmission of the missing data packet. Packet header 310 may also include a time stamp, typically representing an offset time value from the beginning of file 300. By reading the time stamp, which is generally assigned by the video recorder and/or the encoder, the client application may be able to ascertain when video data pertaining to a particular data packet should be rendered.

In one embodiment, the digital video data is encoded into a series of independent (I) frames and dependent (P) frames, which have been encoded in

accordance with a Hierarchical Vector Quantization encoding technique. An implementation of this technique has been discussed in detail in the commonly assigned, co-pending patent application, U.S. Patent Application Serial Number 08/623,299, filed March 28, 1996, which is incorporated herein by reference. In general, an I frame contains all the data necessary for decoding into a video frame without depending on data from any other frame. A P frame, on the other hand, encodes only the difference between the video frame it represents and the most recent I frames. Since P frames only encode the differences, they are generally smaller and therefore more efficient to transmit and store.

In accordance with one aspect of the present invention, the video data pertaining to a given session is encoded in two separate data streams: a digital play stream and a digital fast forward stream. To facilitate discussion, Fig. 4 depicts, in accordance with one embodiment of the invention, two video streams: a play stream 400 and a fast-forward stream 450. It should be borne in mind that although the streams are depicted by a series of video frames in Fig. 4 for ease of discussion, each frame in practice may be transmitted in a portion of a data packet, or in one or more data packets. Accordingly, the fast-forward streams includes a plurality of underlying data packets representing the fast-forward video frames, and the play stream includes a plurality of underlying data packets representing the playable video frames.

In the play mode, data packets associated with play stream 400 are streamed from the server to the client to be rendered by the renderer application. The client computer looks at the packet header associated with each of packets 402-424 and ascertains whether they are playable. Since they are associated with the play stream, the packet header in each will indicate that each of packets 402-424 is playable. The frame rate rendered may be, for example, 10 frames per second in one embodiment but may vary to suit the requirements and capabilities of a particular system.

As mentioned earlier, there is provided, in accordance with one embodiment of the present invention, a separate fast forward stream. The use of the fast forward stream may advantageously lower the bandwidth requirement while the renderer application is in the fast forward mode while improving the quality of the fast forward video frames. To discuss this aspect of the invention, reference should now be made to fast forward stream 450 and compares this fast forward stream to play stream 400.

Before the use of fast forward stream 450 is discussed, consider the situation wherein play stream 400 is employed for fast forwarding instead. In play stream 400, the I frames are also fast forwardable in addition to being playable. If play stream 400 is also employed in fast forwarding, at least the I frames must be transmitted to the

client computer for display. The I frames must be transmitted since P frames cannot be decoded at the client computer unless information pertaining the I frames are available to the decoder. In this example, I frames 401, 410, and 418 would be transmitted. Of course I frame 410 may be omitted when fast forwarding as well but this may  
5 negatively impact the display quality of the fast forward video, e.g., makes the video clip more jerky, since only two frames (I frames 402 and 418) are available for display at the client computer when fast forwarding through this portion of the video clip.

The fast forwarding visual effect is achieved when some frames are discarded and the remaining frames are rendered in a shorter period of time than it would have  
10 taken to render all the corresponding frames in the play stream. In the present example, if P frames 404, 406, 408, 412, 414, and 416 are discarded (i.e., not transmitted from the server computer to the client computer) and I frames 402, 410, and 418 are rendered in a shortened period of time (e.g., 3/9 the time it takes to play frames 402-418 in a play mode), fast forwarding would have been achieved using only the play stream 400.

Compare now to the situation wherein the fast forward stream 450 is employed to achieve the fast forward effect. In Fig. 4, the time stamp of I frame 402 in play stream 400 roughly corresponds to the time stamp of I frame 452 in the fast forward stream 450. Likewise, the time stamp of I frame 418 in play stream 400 roughly corresponds in time to the time stamp of I frame 460 in the fast forward stream 450.  
20 Note that the data contained in the other six P frames (404, 406, 408, 412, 414, and 416) and one other I frame (410) of play stream 400 is encoded in three P frames (456, 458, and 460) of fast forward stream 450. When fast forward stream 450 is employed in the fast forward mode, the same portion of the video clip is rendered in 2 I frames (452 and 460) and 3 P frames (454, 456, and 458). As discussed above, this same  
25 portion of the video clip would have required three I frames (402, 410, and 418) had play stream 400 been employed for fast forwarding. In most cases, it requires less bandwidth to transmit three P frames (454, 456, and 458) than an additional I frame (410) since P frames, as mentioned, only encode the differences. Accordingly, the use of fast forward stream 450 advantageously reduces the bandwidth requirement while  
30 the renderer application is in the fast forward mode.

As an added advantage, the use of fast forward stream 450 permits the same portion of the video clip to be rendered in five frames (452-460) instead of only 3 I frames (402, 410, and 418) as would have been the case had play stream 400 been employed for fast forwarding. As can be appreciated by those skilled, the availability  
35 of two additional frames reduces the jerkiness of the rendered video clip, thereby resulting in a smoother fast forward sequence.

In accordance with one particularly advantageous embodiment of the present invention, there are provided two play modes: live play mode and real-time play mode. Live play mode represents the operating mode of the renderer application when the user at the client computer indicates that he would like to view the video frames as they are being streamed from the encoder (via the server). Typically, live play mode displays data packets substantially as the event they represent is recorded (with some minor delay due to encoding and transmission). In the live play mode, it is typically not permissible to fast forward (since the event to be captured by subsequent video frames has yet to occur).

Real-time play mode, on the other hand, represents the operating mode of the renderer application when the user at the client computer indicates that he would like to view video frames that have been stored earlier in the server, i.e., in a digital video file. The digital video file may have been completely formed or partially formed (e.g., the case where the user rewinds to past video frames of a live video stream). Accordingly, it is possible to fast forward while in real-time play mode since subsequent frames may have already been encoded and pre-stored at the server. It should be pointed out again that in a given live video stream, the user may be in real-time play mode when watching events that have been encoded and pre-stored some time ago. If the user fast forward on the same live stream, he may encounter the end of the pre-stored file (which is still growing as new video frames are captured, encoded, and stored), and may need to switch to a live-play mode to view the more recently recorded and encoded video frames.

Fig. 5 illustrates, in accordance with one embodiment of the invention, a simplified flowchart illustrating the implementation of certain control features such as play and fast forward. Fig. 5 is a flow chart for the general case; specific control features are discussed in greater detail with reference to the general flow chart of Fig. 5 and in the drawings and text hereinbelow. In step 502, the client issues a command at the client computer, e.g., by clicking on an appropriate icon or interacting via a dialog box. In the context of the current discussion, this command may represent for example a play command. The play command typically involves the command itself and a time parameter. For example, the user may wish to play in live play mode, in which case the time parameter may simply represent a flag for the server to stream the latest encoded frames to the client computer for display. As a further example, the user may wish to play from time 2 seconds, i.e., 2 seconds from the start of the video file. If the recording session has been longer than 2 seconds (e.g., the currently recorded and encoded video frame has a time stamp of 5 minutes) the command to play from time 2 seconds indicates that video frames pre-stored at the server need to be recalled for

rendering, i.e., the render application should operate in the real-time play mode but not the live play mode.

In step 504, the server ascertains the first data packet to send in order to satisfy the request of step 502. In the context of the play command, step 504 is implemented differently depending on whether the play mode is real-time or live. Fig. 6 illustrates the steps involved, in one embodiment of the present invention, to implement step 504 for the real-time play mode. Using the above real-time play mode example, the server seeks backward on the play stream, e.g., play stream 400 of Fig. 4, among the seekable frames to the seekable frame that is right before the specified time (in step 532). As mentioned in the above example, the specified time is 2 seconds. If the frame corresponding to 2 seconds is P frame 414, for example, the server seeks backward to I frame 410, i.e., the seekable I frame before P frame 414 (I frames, in addition to being playable and fast-forwardable, are also seekable). I frame 410 will need to be streamed to the client application since it contains the information necessary to decode P frame 414, which is the frame the user wants to start playing from. In this case, the first data packet streamed will be the data packet(s) representing I frame 410 (step 534). Alternatively, the server may seek forward on the play stream to the next seekable frame, which results in I frame 418 being the first video frame streamed out.

Fig. 7, on the other hand, illustrates the steps involved in one embodiment of the present invention in implementing step 504 for the live play mode. In the live play mode, the user wishes to watch from the currently encoded video frames. In this case, the server simply monitors the stream of encoded video frames as they are sent to the server from the encoder. The seekable video frame that arrives next at the server, i.e., the next I frame, is then the first video frame to be sent to the client application since the information in this I-frame may be used to decode itself and subsequent P frames to allow video rendering to take place in real-time (step 552). Accordingly, the first data packet streamed will be the data packet(s) representing this I frame (step 534).

Referring back now to Fig. 5, the server and the client computer may flush their buffers in step 506, e.g., server play-out buffer 112, retransmit buffer 118, and client play-out buffer 120 of Fig. 1. The buffers are flushed since they may already contain data packets that were transmitted and/or received prior to the issuance of the latest command by the user (via the client computer). In the context of the present example, the renderer application may be in the real-time play mode before the user issues a command to play live, and the buffers may already contain therein data packets pertaining to the pre-stored video frames and need to be flushed so that new video frames pertaining to the live video frames may be transmitted and received.

In step 508, the server communicates to the client the first valid packet sequence number of the series of packets to be rendered to satisfy the request of step 502. In step 510, the server streams data packets starting from the data packet ascertained in step 504. The first data packet typically includes the packet sequence number communicated earlier to the client computer in step 508. The client computer monitors the streams of data packets received and ignores all data packets whose sequence numbers may be smaller than the packet sequence number communicated in step 508. This is because some data packets associated with the earlier modes may have been in transit and may just now arrive at the client computer. Since the client computer knows in advance the sequence number of the first data packet it needs, the client computer may simply ignore these subsequently arriving data packets and waits for the proper data packets before rendering. Once the proper data packets arrive at the client computer, the client computer may begin to render them (step 512) starting from the data packet having the sequence number communicated earlier in step 508.

While playing in the pre-stored mode (or any other mode except live play), the user may fast forward to either a subsequent video frame in the pre-stored video file or to the current live video frame. As a side note, fast forward is not possible while the user is in the live play mode since the event to be recorded has yet to occur, much less to be digitized and encoded for transmission. The steps involved in the transition between, for example, real-time play and fast forward may be better understood with reference back to Fig. 5.

If the renderer application is currently in the real-time play mode, a command to fast forward by the user results in a fast forward command, along with a time parameter being sent from the client computer to the server computer (step 502). In one embodiment, the time parameter indicates to the server the time stamp of the most recently rendered video frame to enable to server to ascertain which video frame of the fast forward stream should be sent first in order to assure a smooth transition between play and fast forward. This is because typically only the client knows which video frame has been rendered last prior to the issuance of the fast forward command. The server computer may not have this knowledge since the server may have streamed other video frames which are still in transit or in the buffer and have not yet been rendered by the client computer.

In step 504, the server ascertains the first fast forwardable video frame to send to the client computer responsive to the command issued in step 502. In one embodiment, the play stream is employed for fast forwarding. In this embodiment, the first fast forwardable frame represents the I frame that is immediately prior in time to the video frame corresponding to the time parameter sent in step 502. Alternatively, the I

frame that is immediately later in time to the video frame corresponding to the time parameter sent in step 502 may be employed as the first fast forwardable frame.

In a particularly advantageous embodiment, the fast forward stream is employed for fast forwarding. The steps taken by the server to ascertain the first data packet of the fast forward stream to send when transitioning from the real-time play mode (or other modes except live play) to the fast forward mode are depicted in Fig. 8. In step 802, the server seeks back in the fast forward stream from the video frame whose time corresponds or most closely corresponds to the time parameter sent from the client to the server in step 502 for the first frame to be sent. Alternatively, the server may seek forward in the fast forward stream from the video frame whose time corresponds or most closely corresponds to the time parameter sent from the client to the server in step 502 for the first frame to be sent. The fast forward stream, like the play stream, are stored in the server as the recording session progresses as video files.

To illustrate the steps involved in transitioning from the real-time play mode (or other modes except live play) to the fast forward mode when the fast forward stream is involved, consider the following example. Assume that the client is currently in the real-time play mode, and the last video frame rendered by the client computer is P frame 412 in play stream 400 of Fig. 4. Further assume that this P frame 412 has a time stamp of 6 seconds. In step 504, the server switches to the fast forward stream 450 and seeks backward from the video frame whose time corresponds or most closely corresponds to 6 seconds (step 802). Assume that P frame 456 in fast forward stream 450 has a time stamp that corresponds to or most closely corresponds to 6 seconds, the server will seek backward to the previous seekable frame, i.e., I frame 452 (in fast forward stream 450, all I frames are seekable and forwardable while all P frames are forwardable). Once I frame 452 is ascertained, the data packets representing the video frames of fast forward stream 450 are streamed to the client for rendering as a fast forward video clip (steps 506-514 of Fig. 5). It should be appreciated that in most cases, only one stream (either play or fast forward) is transmitted from the server to the client at any given time.

In the reverse situation, the user may also transition from the fast forward mode to the play mode. Assuming fast forward stream 450 is associated with an event still being recorded, the play mode entered depends on whether fast forwarding has reached the end of fast forward stream 450 (stored in the server) prior to transitioning to play (e.g., prior to the user pushing the play button or icon). If the end of the fast forward stream has not been reached while fast forwarding, the play mode entered will be the real-time play mode (i.e., playing previously recorded and stored video frames versus live video frames). This is the situation wherein the user presses the play button or icon

while fast forwarding to begin playing from, for example, the video frame with a time stamp of 4 minutes when the live video frame has a time stamp of, for example, 6 minutes.

5 The steps involved in the transition between fast forward to real-time play may be better understood with reference back to Fig. 5. If the renderer application is currently in the fast forward mode, a command to play by the user results in a play command, along with a time parameter being sent from the client computer to the server computer (step 502). If the time parameter is earlier in time than the time stamp of the current live video frame, it is understood that the renderer application is transitioning to the real-time play mode. In one embodiment, the time parameter indicates to the server the time stamp of the most recently rendered video frame to enable to server to ascertain which video frame of the play stream 400 should be sent first in order to assure a smooth transition between fast forward and play. This is because only the client typically knows which video frame has been rendered last prior to the issuance of the play command (i.e., since the server may have streamed other video frames which are still in transit and have not yet arrived at the client computer).

10 In step 504, the server ascertains the first video frame to send to the client computer responsive to the command issued in step 502. In one embodiment, the server switches to the play stream and seeks back (or in the alternative forward) in the play stream from the video frame whose time corresponds or most closely corresponds to the time parameter sent from the client to the server in step 502 for the first seekable video frame. The steps taken by the server to ascertain the first data packet of the play stream to send when transitioning from the fast forward mode to the real-time play mode (step 504 of Fig. 5) are depicted, in accordance with an embodiment of the present invention, in Fig. 9.

20 To illustrate the foregoing, assume that the last video frame rendered by the client computer is P frame 458 in fast forward stream 450 of Fig. 4, and this P frame 458 has a time stamp of 36 seconds. In step 504, the server switches to the play stream 400 and seeks backward from the video frame whose time corresponds or most closely corresponds to 36 seconds (step 902). Assume that P frame 416 of play stream 400 has a time stamp that most closely corresponds to 36 seconds, the server will seek backward to the previous seekable frame in the play stream (step 904), i.e., I frame 410. Once I frame 410 is ascertained, the data packets representing the video frames of play stream 400 are streamed to the client for rendering as a real-time video clip (steps 506-514 of Fig. 5). If the server seeks forward in the alternative, I frame 418 represents the first video frame to be streamed to the client.



As mentioned earlier, it is not possible to transition from the live play mode to the fast forward mode. However, the user may transition from the fast forward mode to the live play mode (either directly by indicating the desire to the client computer or by fast forwarding to the end of the fast forward file in the server and running out of pre-stored fast forwardable frames). In one embodiment, the transition between the fast forward mode and the live play mode occurs automatically (i.e., without further intervention of the user and/or the client computer) as the end of the file of stored video frames is reached. The steps involved in the transition between fast forward to live play may be better understood with reference back to Fig. 5. If the renderer application is currently in the fast forward mode, a command to play live by the user results in a play command, along with a time parameter being sent from the client computer to the server computer (step 502). In one embodiment, the time parameter may represent a flag that simply indicates to the server that the client would like to play live video.

In step 504, the server ascertains in the play stream the first video frame to send to the client computer responsive to the command issued in step 502. In one embodiment, the server simply waits until the next seekable video frame of the play stream 400 is received. This seekable video frame (typically an I frame) and subsequent video frames from play stream 400 are then sent from the server to the client for displaying the live video.

While in the live play mode, the user may also enter the rewind mode, e.g., to perform an instant replay of a particularly interesting sequence or to go back and start displaying again from a previously rendered video frame. This is a particularly advantageous aspect of the invention since rewind is enabled, in accordance with one aspect of the present invention, in an efficient manner on a live video stream. In rewinding, the I frames of the play stream are transmitted to the client in reverse chronological order starting from the I frame immediately prior to the video frame whose time stamp is specified by the client.

The steps involved in the transition between live play (or any other mode) to rewind may be better understood with reference back to Fig. 5. If the renderer application is currently in the live play mode (or any other mode), a command to rewind by the user results in a rewind command, along with a time parameter being sent from the client computer to the server computer (step 502). In one embodiment, the time parameter indicates to the server the time stamp of the most recently rendered video frame to enable to server to ascertain which video frame of the play stream 400 should be sent in order to assure a smooth transition between play and rewind. This is because only the client typically knows which video frame has been rendered last prior to the

issuance of the rewind command (i.e., since the server may have streamed other video frames which have not yet arrived at the client computer).

In step 504, the server ascertains the first video frame to send to the client computer responsive to the command issued in step 502. In one embodiment, the server seeks back in the play stream for the first seekable frame (I frame) from the video frame whose time corresponds or most closely corresponds to the time parameter sent from the client to the server in step 502. Alternatively, the server may seek forward in the play stream for the first seekable frame (I frame) from the video frame whose time corresponds or most closely corresponds to the time parameter sent from the client to the server in step 502. It should be appreciated that if the previous mode is fast forward, the server may switch from fast forward stream 450 to play stream 400 prior to ascertaining the first video frame to send to the client to accomplish rewind.

Once the appropriate I frame is ascertained, the server prepares the client for receiving the rewound video frames (steps 506 and 508 of Fig. 5). Thereafter, the server sends I frames of play stream 400 in reverse order (by repetitively seeking back for the previous seekable frame in play stream 400 and sending each to the client, for example). The sending of the I frames in reverse and displaying each at the client take place in steps 510 and 512 of Fig. 5. It should be apparent that the above discussion may also apply when transitioning from the real-time play mode to the rewind mode, or from any mode to the rewind mode, for that matter.

The transition from the rewind mode to the play mode or the fast forward mode follows the steps of Fig. 5, which have been discussed in detail herein. In general, the client sends a rewind command, along with a time parameter to indicate the last video frame that has been rendered by the client while rewinding (step 504). In the rewind case, the last video frame rendered represents one of the seekable frames (i.e., I frames) that were sent to the client in reverse order to facilitate rewinding. Thereafter the server streams data packets (either of the play stream or the fast forward stream depending on which mode the client transitions to) to the client (steps 504-512) starting from the first seekable frame closest to the frame whose time corresponds or most closely corresponds to the time stamp sent..

In accordance with one embodiment of the present invention, the client may enter the stop mode from any other mode. The transition to the stop mode involves sending the server a signal to terminate the current mode and reset the play stream to the first video frame of the video file (e.g., by sending it a time parameter of zero). The transition from the stop mode to any other mode essentially involves the steps of Fig. 5, the specifics of which depend on the mode being transitioned to. For example, the user

may, via the client, fast forward, play beginning from a specific video frame (by first fast forwarding or seeking from the current start of file position), or the like.

In accordance with one embodiment of the present invention, the client may enter the pause mode from any other mode. Pause mode involves terminating display immediately at the client and requesting the server to stop sending new data packets. In one particularly advantageous embodiment, the client sends a signal to the server to set the outgoing bandwidth to zero, thereby terminating the transmission of data packets from the server to the client. Note that unlike the steps of Fig. 5, however, pause does not involve flushing the buffers. In other words, the data packets already sent, received, or are currently in transit through the network are kept and/or continued to be received in the client play-out buffer 120 at the client computer, thereby permitting the renderer application to immediately start rendering video frames from where it left off when the unpause mode is entered.

Upon entering the unpause mode, the client computer may send a signal to the server computer to reset the bandwidth to a nonzero value to permit data packets to be transmitted from where transmission was suspended (upon entering the pause mode). In this manner, the client computer may advantageously immediately begin to render video frames after unpausing using frames buffered at the client play-out buffer associated with the client computer. Note that no latency is suffered upon entering the unpause mode since video frames already exist in the client play-out buffer and were not flushed. Further, synchronization issues are substantially simplified since the server computer simply commences transmission of data packets from it left off when the outgoing bandwidth was reset to zero. Further, the inventive pause and unpause technique maximizes smoothness of the displayed video frames since no frames are missed upon unpausing (as would be the case if the buffers were flushed and the server must begin streaming from the closest seekable I frame).

In accordance with yet another aspect of the present invention, the data packets streamed from the server may be stored, i.e., recorded, at the client computer for subsequent display without having to be retransmitted from the server again. Responsive to a record button or icon, for example, the client computer may record video frames either while displaying the video frames (in either live play or real-time play modes) or by simply requesting the file of video frames to be streamed from the server without displaying the frames at the client computer. Some client applications may also be provided with mechanisms to prevent unauthorized recording or to keep track of recording for administrative, e.g., billing, purposes. It should be appreciated that mechanisms to inhibit unauthorized recording may also be provided at the server computer or even with the data packets themselves.

In another aspect of the present invention, the fast forward stream may be employed to facilitate low-bandwidth play (either live or real-time) when the network's available bandwidth is insufficient to support the transmission of the video frames in the play stream. As can be seen from Fig. 4 and the discussion herein, the fast forward stream encodes video data pertaining to a given video event using fewer frames than the number of frames required by the play stream. When the network's available bandwidth degrades, the fast forward stream may be sent by the server instead of the play stream. The renderer application may then display the frames of the fast forward stream at a slower frame rate than the frame rate normally employed on the fast forward stream's frames to allow the small number of frames of the fast forward stream to be displayed in the same amount of time that the more numerous frames of the play stream would have taken to display (had the play stream been transmitted). Although the resulting video may be more jerky, this aspect of the invention permits video streaming to continue even with very limited network bandwidth, which prevents the timely transmission of the frames in the play stream and would have resulted in termination of play otherwise.

While this invention has been described in terms of several preferred embodiments, there are alterations, permutations, and equivalents which fall within the scope of this invention. For example, although only a subset of all permutations of control mode transitions are discussed in detail herein for brevity's sake, it should be apparent to those skilled in the art given this disclosure that the transitions involving other modes may be practiced in accordance with the general technique disclosed herein. Further, although I frames and P frames are discussed throughout this disclosure, it should be borne in mind that the invention applies for any encoding schemes wherein the video data is encoded as a series of independent and dependent video frames. As a further example, although the client-server computer architecture is employed herein to facilitate ease of discussion, the term "client" denotes any type of computer capable of receiving data from another computer. Likewise, the term "server" denotes any type of computer capable of streaming data to another computer. The above is true irrespective whether the computers are coupled together using what is popularly known as "client-server" architectures. It should also be noted that there are many alternative ways of implementing the methods and apparatuses of the present invention. It is therefore intended that the following appended claims be interpreted as including all such alterations, permutations, and equivalents as fall within the true spirit and scope of the present invention.

**What is claimed is:**

1. A method for transmitting streamed digital video data from a server, said server being configured for coupling to a client computer via a computer network,  
5 comprising:

inputting a first plurality of data packets into a server play-out buffer of said server, said first plurality of said data packets containing video frames representing said streamed digital video data, an output of said server play-out buffer being configured to be coupled to a network data connection for transmitting said first plurality of said data  
10 packets to said client computer;

receiving, using a retransmission buffer, said first plurality of said data packets from said server play-out buffer, an output of said retransmission buffer being coupled to said network data connection; and

outputting said first plurality of said data packets from said server play-out  
15 buffer onto said network data connection for transmitting said data packets to said client computer via said computer network.

2. The method of claim 1 further comprising:

outputting a selected data packet from said retransmission packet to said  
20 network data connection responsive to a retransmission request from said client computer, said selected data packet representing a data packet earlier outputted by said server play-out buffer onto said network data connection but unsatisfactorily received by said client computer.

25 3. The method of claim 2 further comprising:

discarding, responsive to said receiving said first plurality of said data packets at said retransmission buffer, old data packets from said retransmission buffer, said old data packets representing data packets buffered before said first plurality of data packets is inputted into said retransmission buffer.

30

4. The method of claim 3 wherein said retransmission buffer is sized to enable a given data packet to remain in said retransmission buffer for a predefined period of time before being discarded, said predefined time period being longer than an average acceptable latency period between the time a data packet is transmitted from said server computer and the time said data packet is rendered at a renderer application associated with said client computer.

5. The method of claim 1 wherein said video frames represents playable video frames of said streamed digital video data, said playable frames representing frames configured to be displayed while a renderer application associated with said client computer is in the play mode.

6. The method of claim 5 wherein said streamed digital video data represents live video and said play mode represents live play mode.

7. The method of claim 6 wherein said streamed digital video data represents real-time video and said play mode represents real-time video mode.

8. The method of claim 1 wherein said streamed digital video data represents a fast forward stream of video frames, said fast forward stream of video frames represents a first stream of video frames encoded differently from a second stream of video frames, said first stream of video frames being configured for normal play by a renderer application associated with said client computer, both said first stream and said second stream pertain to a recorded event, said first stream of video frames, when streamed to said client computer, permits the user to view said recorded event at a faster pace than a pace afforded by said normal play.

9. A method for displaying live streamed digital video data on a client computer, said client computer being configured to receive said live streamed digital video data from a server computer via a computer network, said live streamed digital video data being transmitted from said server computer to said client computer as a stream of video frames, comprising:

receiving a first plurality of video frames at said client computer, said plurality of video frames representing a subset of said stream of video frames, said stream of video frames comprising independent playable video frames and dependent playable video frames;

5 displaying said first plurality of video frames on a video display terminal associated with said client computer; and

issuing a rewind command from said client computer to said server, said rewind command causing a second plurality of video frames of said stream of video frames different from said first plurality of video frames to be streamed from said server  
10 computer to said client computer, said second plurality of video frames has been streamed at least once to said client computer.

10. The method of claim 9 further comprising flushing a client play-out buffer associated with said client computer to prepare said client play-out buffer to  
15 receive said second plurality of video frames, said client play-out buffer representing a buffer employed to buffer video frames received from said server computer prior to said displaying.

11. The method of claim 9 wherein said second plurality of video frames  
20 represents rewound independent playable video frames, said rewound independent playable video frames representing independent playable video frames of said stream of video frames that are prior in time to a video frame rendered immediately prior to said issuing said rewind command and that are sent in reverse chronological order from said server computer to said client computer.

25

12. The method of claim 11 wherein said issuing said rewind command further comprising:

issuing from said client computer to said server computer a time parameter, said time parameter representing a time stamp associated with said video frame rendered  
30 immediately prior to said issuing said rewind command.

13. The method of claim 12 wherein said time parameter is employed by said server to ascertain said video frame rendered immediately prior to said issuing said rewind command in a digital video data file, said digital video data file containing copies of said live streamed digital video data.

5

14. The method of claim 13 wherein a first independent playable video frame of said rewound independent playable video frames is ascertained by seeking backward in said digital video data file from said video frame rendered immediately prior to said issuing said rewind command for an immediately prior in time independent  
10 playable video frame.

15. The method of claim 13 wherein a first independent playable video frame of said rewound independent playable video frames is ascertained by seeking forward in said digital video data file from said video frame rendered immediately prior  
15 to said issuing said rewind command for an immediately later in time independent playable video frame.

16. The method of claim 9 further comprising:

issuing a real-time play command from said client computer after said issuing  
20 said rewind command, said issuing said real-time play command causing a third plurality of video frames to be streamed from said server computer to said client computer for display on said video display terminal, said third plurality of video frames representing video frames that are later in time to a last displayed video frame of said second plurality of said video frames prior to said issuing said real-time play command.

25

17. The method of claim 16 wherein said issuing said real-time play command further comprising:

issuing a time parameter associated with said last displayed video frame of said second plurality of said video frames, said time parameter is employed by said server  
30 computer to ascertain said last displayed video frame of said second plurality of said video frames in a digital video data file, said digital video data file containing copies of said live streamed digital video data.



18. The method of claim 17 wherein a first video frame of said third plurality of said video frames is ascertained by seeking backward in said digital video data file from said last displayed video frame for an immediately prior in time  
5 independent playable video frame.

19. The method of claim 17 wherein a first video frame of said third plurality of said video frames is ascertained by seeking forward in said digital video data file from said last displayed video frame for an immediately later in time  
10 independent playable video frame.

20. In a computer network having a server computer coupled to a client computer, a method for implementing control features configured for controlling a display of streamed digital video data as said client computer transitions from a first  
15 control mode to a second control mode, said streamed digital video data being transmitted from said server computer to said client computer as a stream of video frames comprising independent video frames and dependent video frames, said stream of video frames being encapsulated in a plurality of data packets each having a unique packet sequence number and a unique timestamp, comprising:

20 sending a control command and a time parameter from said client computer to said server computer, said control command representing a command to said server to transmit a first plurality of video frames of said stream of video frames to said client computer in accordance with said second control mode;

25 ascertaining, responsive to said control command and using said server computer, a first independent video frame to transmit to said client computer, said first independent video frame being selected responsive to said time parameter;

transmitting from said server computer to said client computer a packet sequence number associated with said first independent video frame;

30 streaming said first plurality of video frames of said stream of video frames starting from said first independent video frame from said server computer to said client computer to permit said first plurality of video frames to be displayed at said client computer.

21. The method of claim 20 wherein said first control mode is a stop mode and said second control mode is a real-time play mode, said real-time play mode representing a mode wherein said client computer displays video frames starting from a time specified by said time parameter.

22. The method of claim 20 wherein said first control mode is a stop mode and said second control mode is a live play mode, said live play mode representing a mode wherein said client computer displays live video frames as they are received from an encoder coupled to said server computer.

23. The method of claim 20 wherein said first control mode is a live play mode and said second control mode is a rewind mode, said live play mode representing a mode wherein said client computer displays live video frames as they are received from an encoder coupled to said server computer, said rewind mode representing a mode wherein said client computer displays independent video frames in reverse chronological order starting from an independent video frame whose time stamp most closely corresponds to said time parameter.

24. The method of claim 20 wherein said first control mode is a real-time play mode and said second control mode is a fast forward mode, said real-time play mode representing a mode wherein said client computer displays pre-stored video frames, said fast forward mode representing a mode wherein said client computer displays forwardable video frames starting from an independent video frame whose time stamp most closely corresponds to said time parameter, said forwardable video frames being configured to permit a recorded event to be displayed at a faster pace than a pace of said real-time play mode.

25. The method of claim 24 wherein said forwardable video frames represent video frames of a fast forward stream of video frames, said fast forward stream of video frames representing a stream of video frames different from a play stream of video frames employed for said real-time play, said fast forward stream and said play stream being encoded from a single recorded event.

26. The method of claim 20 wherein said first control mode is a live play mode and said second mode is a pause mode, said live play mode representing a mode wherein said client computer displays live video frames as they are received from said server computer, said pause mode representing a mode wherein said stream of video frames is transmitted to said client computer at a bandwidth of zero.

27. The method of claim 20 wherein said first control mode is a fast forward mode and said second mode is a live play mode, said live play mode representing a mode wherein said client computer displays live video frames as they are received from an encoder coupled to said server computer, said fast forward mode representing a mode wherein said client computer displays forwardable video frames of a stored video stream starting from an independent video frame whose time stamp most closely corresponds to said time parameter, said forwardable video frames being configured to permit a recorded event to be displayed at a faster pace than a pace of said live play mode, said first controlled mode being transitioned to said second control mode automatically without intervention from an user of said client computer as an end of said stored video stream is reached.

28. A computer readable medium containing computer-readable instructions for implementing control features configured for controlling a display of streamed digital video data at a client computer as said client computer transitions from a first control mode to a second control mode, said client computer being configured for coupling to a server computer via a computer network, said streamed digital video data being transmitted from said server computer to said client computer as a stream of video frames comprising independent video frames and dependent video frames, said stream of video frames being encapsulated in a plurality of data packets each having a unique packet sequence number and a unique timestamp, said computer readable instructions comprising:

computer readable instructions for sending a control command and a time parameter from said client computer to said server computer, said control command representing a command to said server to transmit a first plurality of video frames of said stream of video frames to said client computer in accordance with said second control mode;

computer readable instructions for ascertaining, responsive to said control command and using said server computer, a first independent video frame to transmit to said client computer, said first independent video frame being selected responsive to said time parameter;

- 5            computer readable instructions for transmitting from said server computer to said client computer a packet sequence number associated with said first independent video frame;

- 10           computer readable instructions for streaming said first plurality of video frames of said stream of video frames starting from said first independent video frame from said server computer to said client computer to permit said first plurality of video frames to be displayed at said client computer.

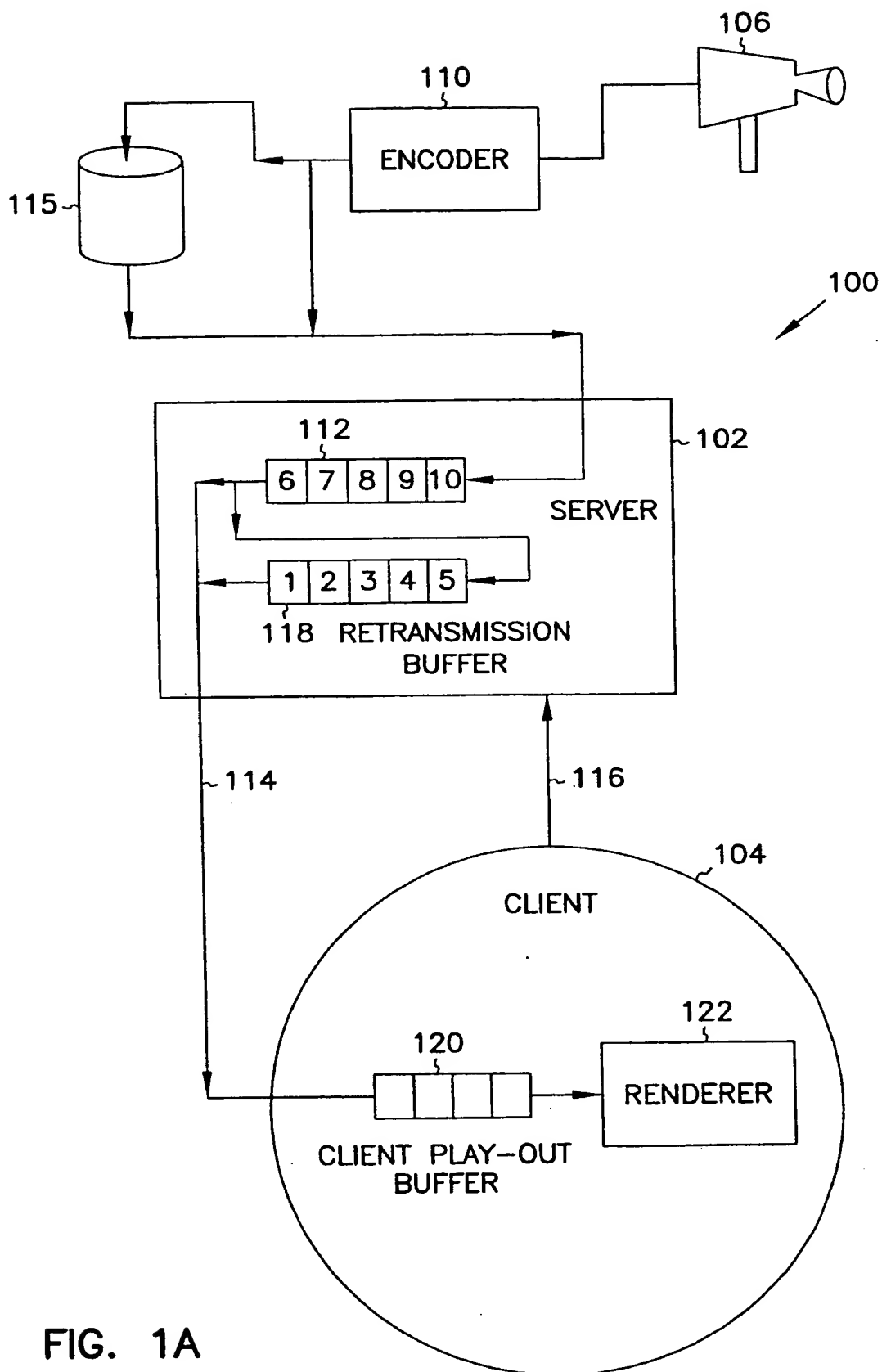
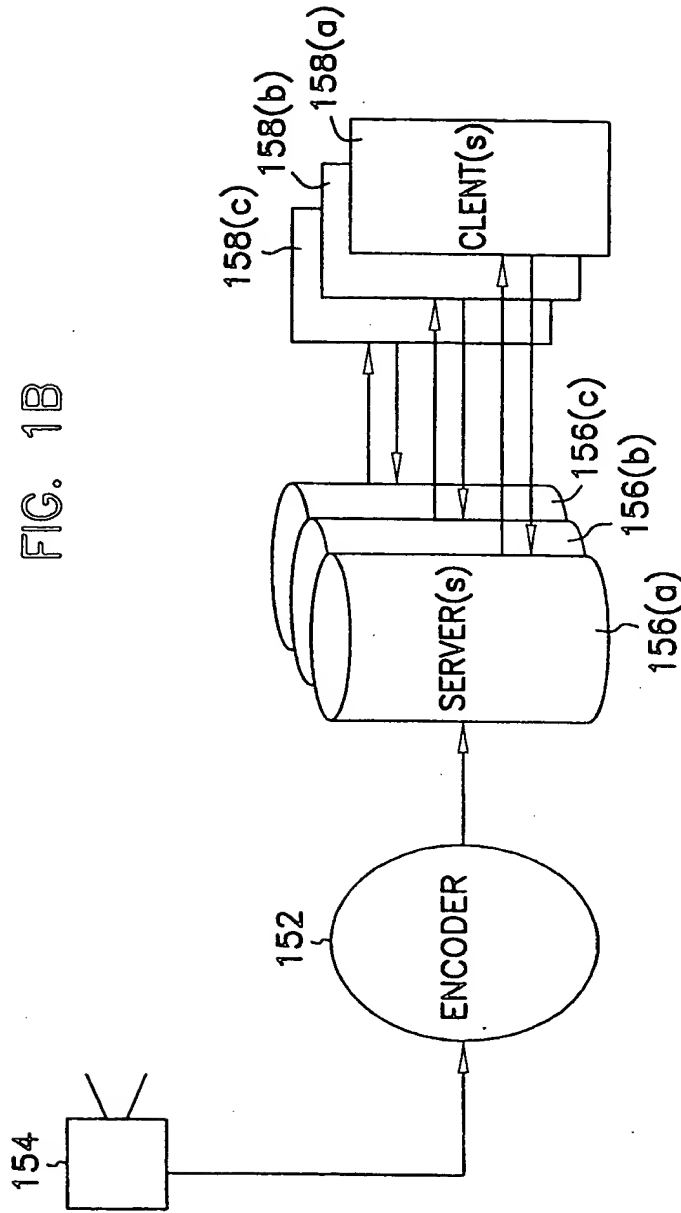


FIG. 1A



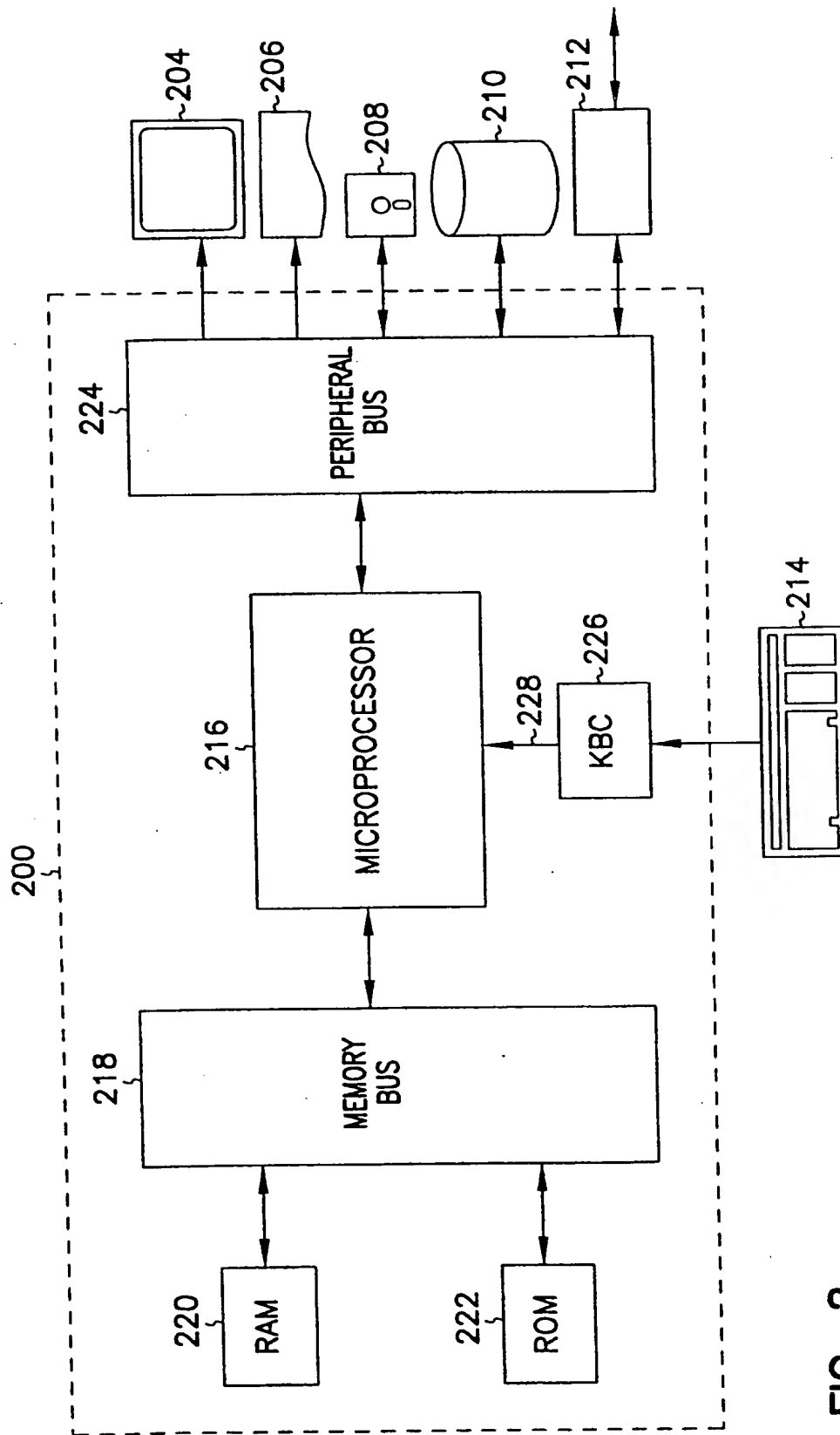


FIG. 2

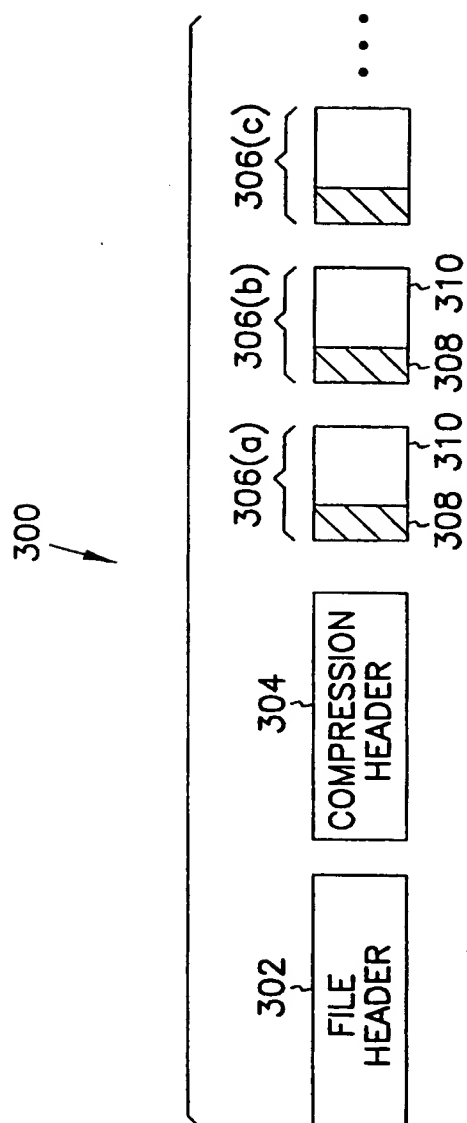


FIG. 3



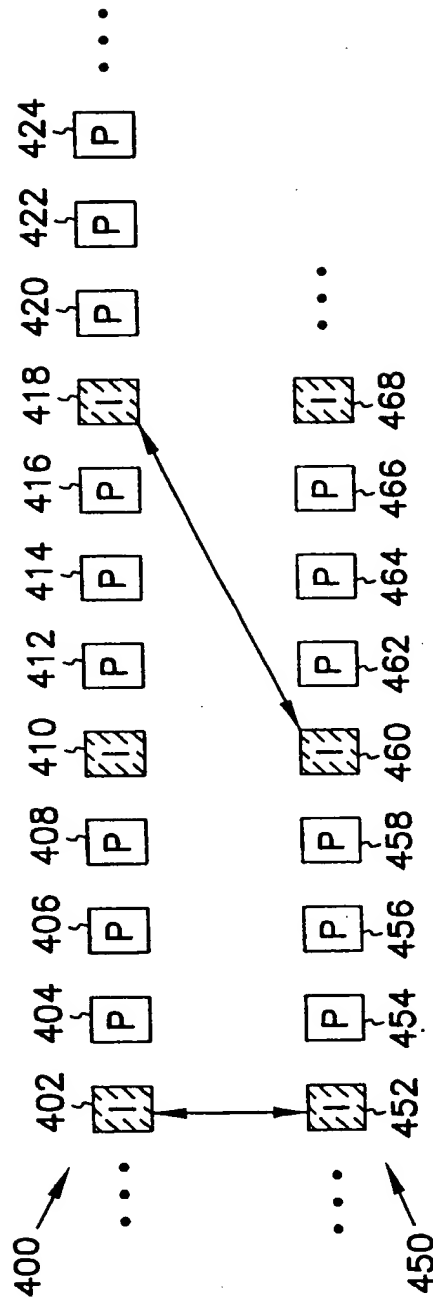


FIG. 4

6/8

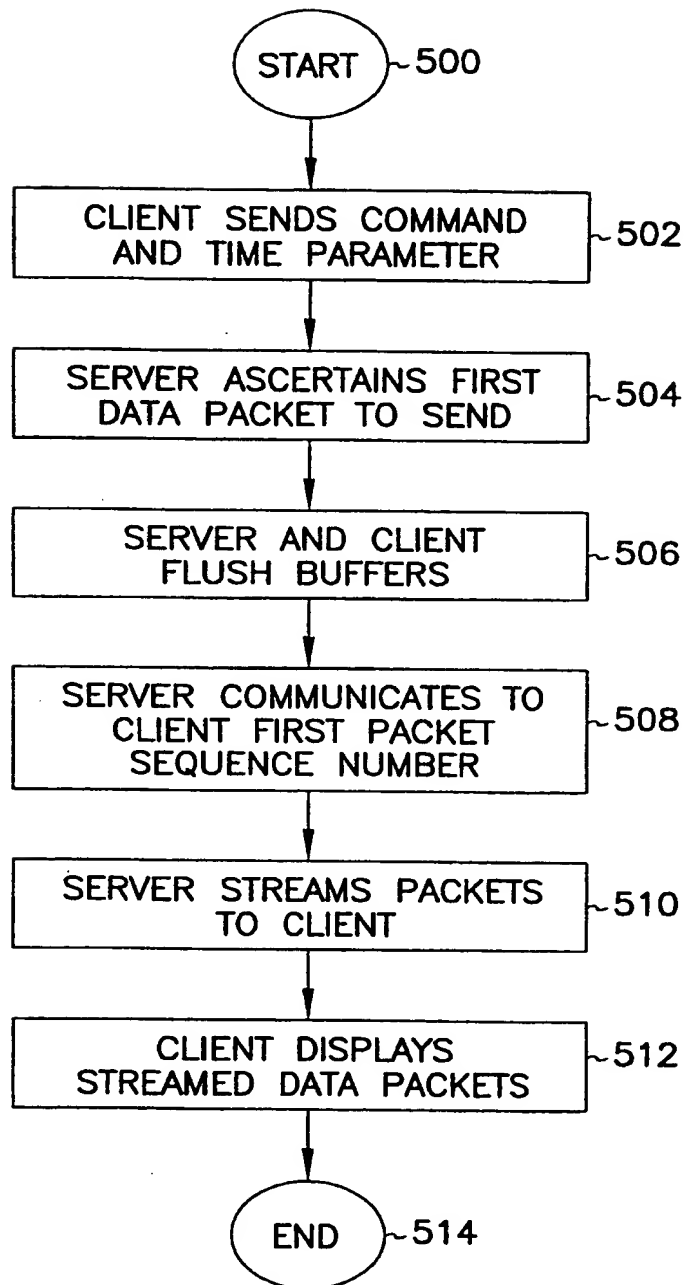
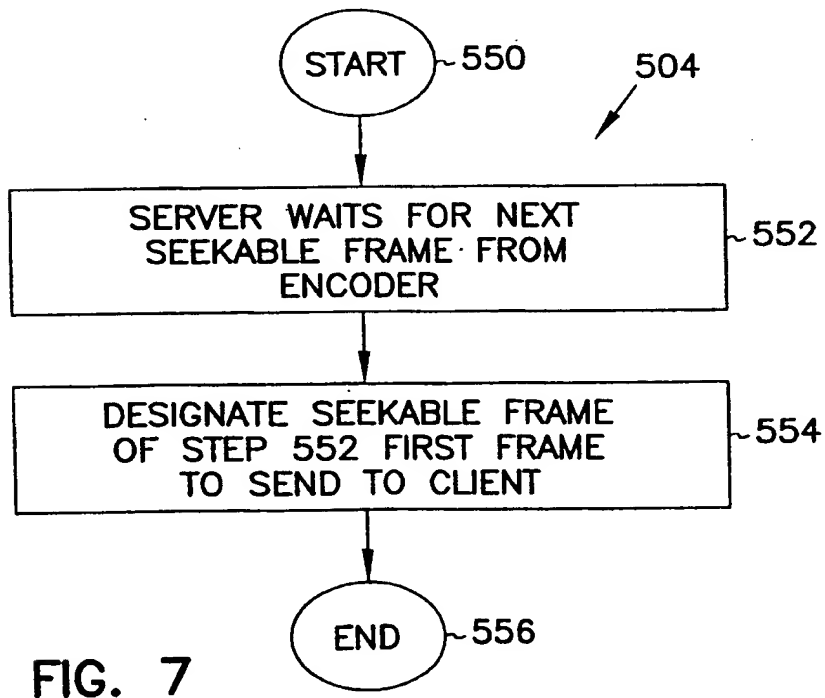
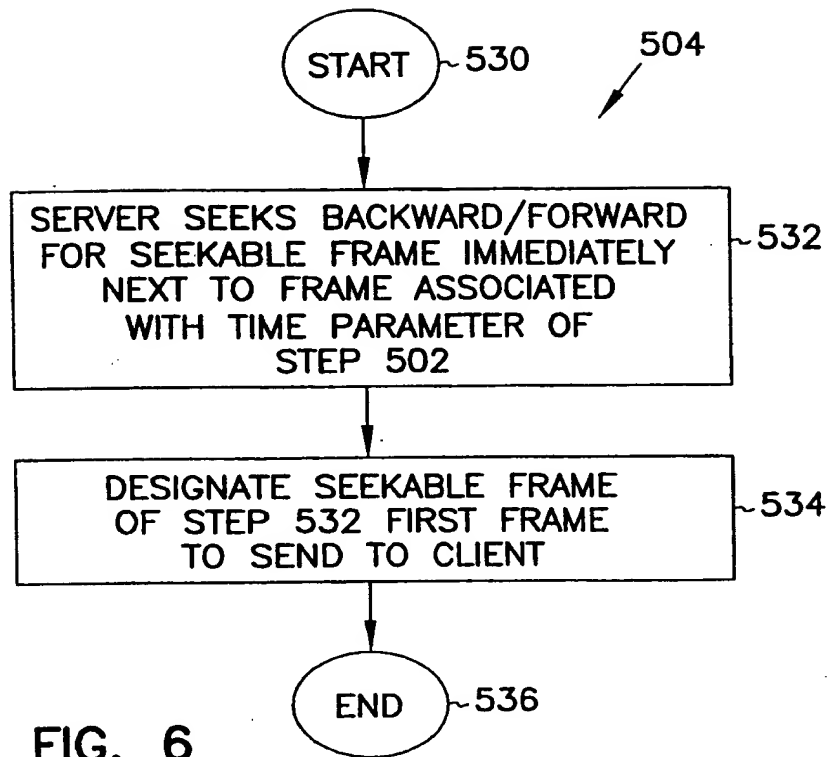
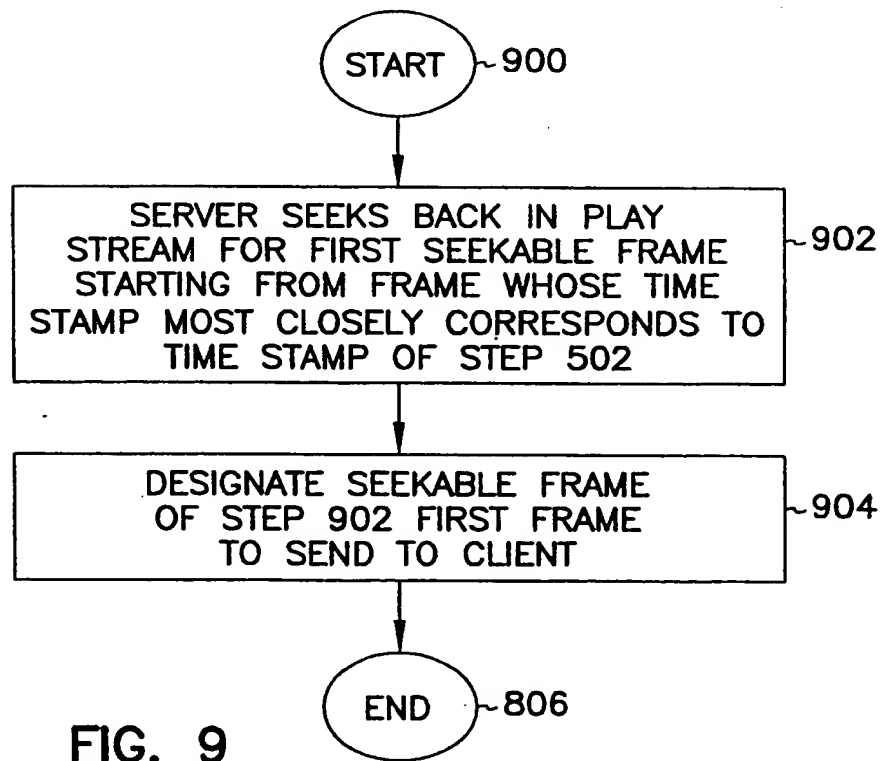
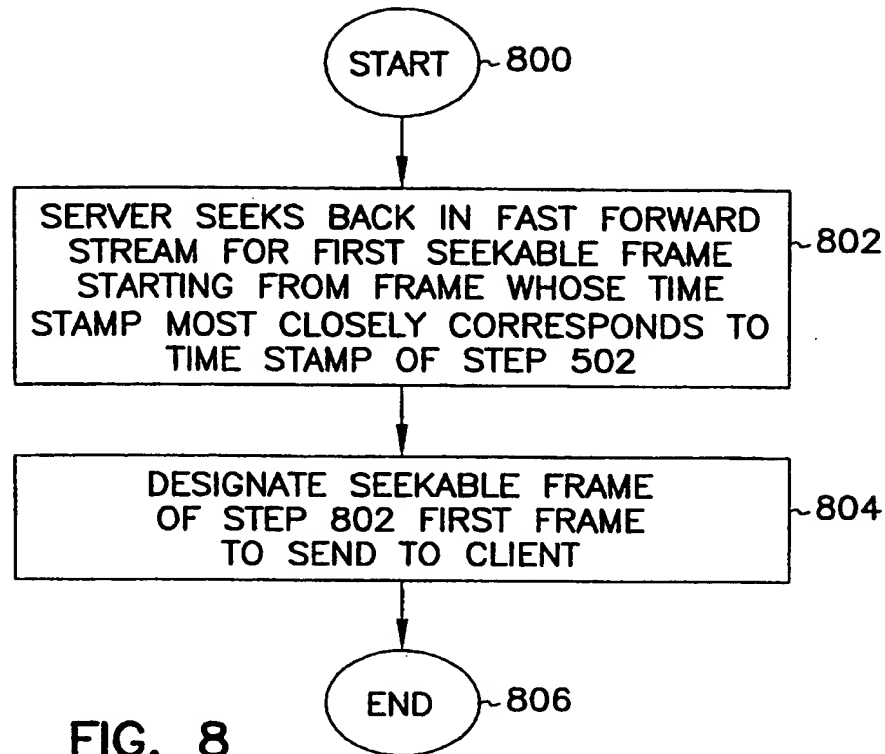


FIG. 5

7/8



8/8



# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 98/01810

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6 H04N7/173

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 5 414 455 A (HOOPER DONALD F ET AL) 9 May 1995 see column 1, line 60 - column 2, line 48  see column 3, line 8 - line 46; figure 1 see column 4, line 34 - line 58; figure 2 see column 7, line 63 - column 8, line 43; figure 4 ---	1  2,5, 8-10,16, 20-24, 26,27
A	EP 0 676 898 A (IBM) 11 October 1995  see the whole document ---	1,2,5,9, 10,20
	-/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

11 May 1998

Date of mailing of the international search report

20/05/1998

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Fuchs, P

## INTERNATIONAL SEARCH REPORT

Intern. Application No

PCT/US 98/01810

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 605 115 A (AT & T CORP) 6 July 1994  see column 1, line 50 - column 3, line 19 see column 18, line 13 - column 19, line 38; figures 9,10 ---	1,5,6, 8-11,16, 20-24, 26,27
A	US 5 341 474 A (GELMAN ALEXANDER ET AL) 23 August 1994  see column 9, line 50 - column 12, line 32; figures 3,4,6,7A-7F ---	1,5,8,9, 11,16, 20-24, 26,27
A	EP 0 653 884 A (BELL TELEPHONE MFG) 17 May 1995  see column 2, line 38 - column 5, line 45 ---	1,5,8,9, 20-24, 26,27
A	CHEN H J ET AL: "A SCALABLE VIDEO-ON-DEMAND SERVICE FOR THE PROVISION OF VCR-LIKE FUNCTIONS" 15 May 1995, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MULTIMEDIA COMPUTING AND SYSTEMS, WASHINGTON, MAY 15 - 18, 1995, PAGE(S) 65 - 72, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS XP000632092 see the whole document ---	1,5,8,9, 20-24, 26,27
A	EP 0 746 158 A (IBM) 4 December 1996 see column 4, line 45 - column 5, line 57 see column 10, line 34 - column 12, line 39; figure 1 -----	28

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Internal Application No

PCT/US 98/01810

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 5414455	A	09-05-1995	CA	2127347 A	08-01-1995
			EP	0633694 A	11-01-1995
EP 0676898	A	11-10-1995	US	5521630 A	28-05-1996
			AU	674621 B	02-01-1997
			AU	8180394 A	12-10-1995
			BR	9501428 A	07-11-1995
			CA	2142801 A	05-10-1995
			CN	1122480 A	15-05-1996
			JP	7284042 A	27-10-1995
			SG	28234 A	01-04-1996
EP 0605115	A	06-07-1994	US	5442389 A	15-08-1995
			JP	7177492 A	14-07-1995
US 5341474	A	23-08-1994	WO	9521414 A	10-08-1995
			EP	0742920 A	20-11-1996
			JP	9502071 T	25-02-1997
			NO	963090 A	01-10-1996
EP 0653884	A	17-05-1995	AU	7774294 A	25-05-1995
			CA	2135990 A	18-05-1995
			CN	1110456 A	18-10-1995
			JP	7203418 A	04-08-1995
			NZ	264831 A	26-11-1996
EP 0746158	A	04-12-1996	US	5646676 A	08-07-1997
			JP	9065289 A	07-03-1997

**THIS PAGE BLANK (USPTO)**